

**PARTIAL LEAST SQUARES AND NEURAL NETWORKS FOR QUANTITATIVE CALIBRATION OF LASER-INDUCED BREAKDOWN SPECTROSCOPY (LIBS) OF GEOLOGIC SAMPLES.** R. B. Anderson<sup>1</sup>, R.V. Morris<sup>2</sup>, S.M. Clegg<sup>3</sup>, S. D. Humphries<sup>3</sup>, R. C. Wiens<sup>3</sup>, J.F. Bell III<sup>1</sup>, and S. A. Mertzman<sup>4</sup>. <sup>1</sup>Cornell University, Department of Astronomy, Ithaca, NY 14850 (randerson@astro.cornell.edu), <sup>2</sup>NASA Johnson Space Center, 2101 NASA Parkway, Houston, TX 77058, <sup>3</sup>Los Alamos National Laboratory, P.O. Box 1663 MS J565, Los Alamos, NM 87545, <sup>4</sup>Dept. of Geosciences, Franklin and Marshall College, Lancaster, PA 17604.

**Introduction:** The ChemCam instrument [1] on the Mars Science Laboratory (MSL) rover will be used to obtain the chemical composition of surface targets within 7 m of the rover using Laser Induced Breakdown Spectroscopy (LIBS). ChemCam analyzes atomic emission spectra (240-800 nm) from a plasma created by a pulsed Nd:KGW 1067 nm laser. The LIBS spectra can be used in a semiquantitative way to rapidly classify targets (e.g., basalt, andesite, carbonate, sulfate, etc.) and in a quantitative way to estimate their major and minor element chemical compositions.

Quantitative chemical analysis from LIBS spectra is complicated by a number of factors, including chemical matrix effects [2]. Recent work has shown promising results using multivariate techniques such as partial least squares (PLS) regression and artificial neural networks (ANN) to predict elemental abundances in samples [e.g. 2-6]. To develop, refine, and evaluate analysis schemes for LIBS spectra of geologic materials, we collected spectra of a diverse set of well-characterized natural geologic samples and are comparing the predictive abilities of PLS, cascade correlation ANN (CC-ANN) and multilayer perceptron ANN (MLP-ANN) analysis procedures.

**Methods and Samples:** LIBS spectra were acquired with a laboratory simulation of the MSL ChemCam instrument. It uses a Nd:YAG laser operating at 1064 nm, pulsed at 10 Hz with a pulse energy of 17 mJ/pulse. Targets were placed in a vacuum chamber filled with ~7 Torr CO<sub>2</sub> at a distance of 7 m. Plasma emission was collected with a telescope and directed through a fiber into one of three Ocean Optics HR2000 spectrometers covering the UV (245.00-325.97 nm), VIS (381.86-471.03 nm) and VNIR (494.93-927.06 nm) wavelength regions.

We selected a diverse suite of rock slab samples from the Mars analog sample collection at NASA's Johnson Space Center [7]. Their chemical compositions are known from independent XRF analysis. This sample collection has been used to validate the calibration of previous instruments on orbital and landed Mars missions including CRISM, OMEGA, MER Pancam, Mini-TES, and Mössbauer, and Phoenix SSI. Our ChemCam samples are common to those being analyzed (in different physical forms) by testbed versions

of the MSL CheMin and SAM instruments and by the flight Mastcam instrument.

Typically, the rock slabs were analyzed 5 to 7 at time along with 5 pressed powder geostandards [8]. When possible, we chose standards with a similar composition to the geologic samples. The geostandards included andesites, basalts, dolomites, gypsum and olivine. LIBS spectra were acquired for 5 spots on each rock slab and geostandard target.

**Data Analysis:** As a first step in data analysis, the UV, VIS and VNIR spectra for each point on each target were normalized and combined to form a single spectrum with 6144 channels. The intensity of atomic emission lines correspond to atomic fraction rather than oxide weight percent, so all standard and sample composition values were converted to atomic fractions. In our analysis for both PLS and ANNs, half of the geologic samples were used to create models for the major rock-forming elements which were then used to predict atomic fractions in the remaining samples based on their spectra (Figure 1).

*Partial least squares:* We used the commercial Unscrambler program [9] to conduct PLS analysis. The PLS algorithm [10] identifies wavelengths in the spectrum for which the intensity correlates (or anticorrelates) with the element of interest, resulting in a set of one weight per wavelength in the spectrum. PLS can be used to model a single variable (PLS1) or multiple variables (PLS2). The results shown here are based on PLS2 models for the major rock-forming elements.

*Artificial neural networks:* ANNs have also been used for predicting elemental abundances based on LIBS spectra. [4-5] ANNs are inherently nonlinear and have a high tolerance for noise, both of which are desirable for analyzing remote LIBS data. ANNs are based on numerous, interconnected simple processing elements ("neurons"), organized as layers with the output of one layer serving as the input for the following layer(s). Previous work using ANNs for LIBS analysis used MLP-ANNs with an input layer, one intermediate "hidden" layer and an output layer [4-5]. One of the drawbacks of MLP-ANNs is that the optimum number of neurons in the hidden layer(s) can only be determined by trial and error. It is prohibitively time consuming to conduct this trial and error method with the entire LIBS spectrum, so we have taken a similar ap-

proach to Sirven et al. [4] and selected several strong lines for each element of interest and trained MLP-ANNs based on the intensity of emission at those wavelengths only.

We also used the cascade correlation (CC) algorithm to train ANNs [10]. CC begins with only an input and output layer and adds new intermediate neurons one at a time. CC-ANNs learn rapidly and determine their own size and topology. Because of the learning speed and simplicity of the networks generated with CC, it was feasible to input the entire LIBS spectrum.

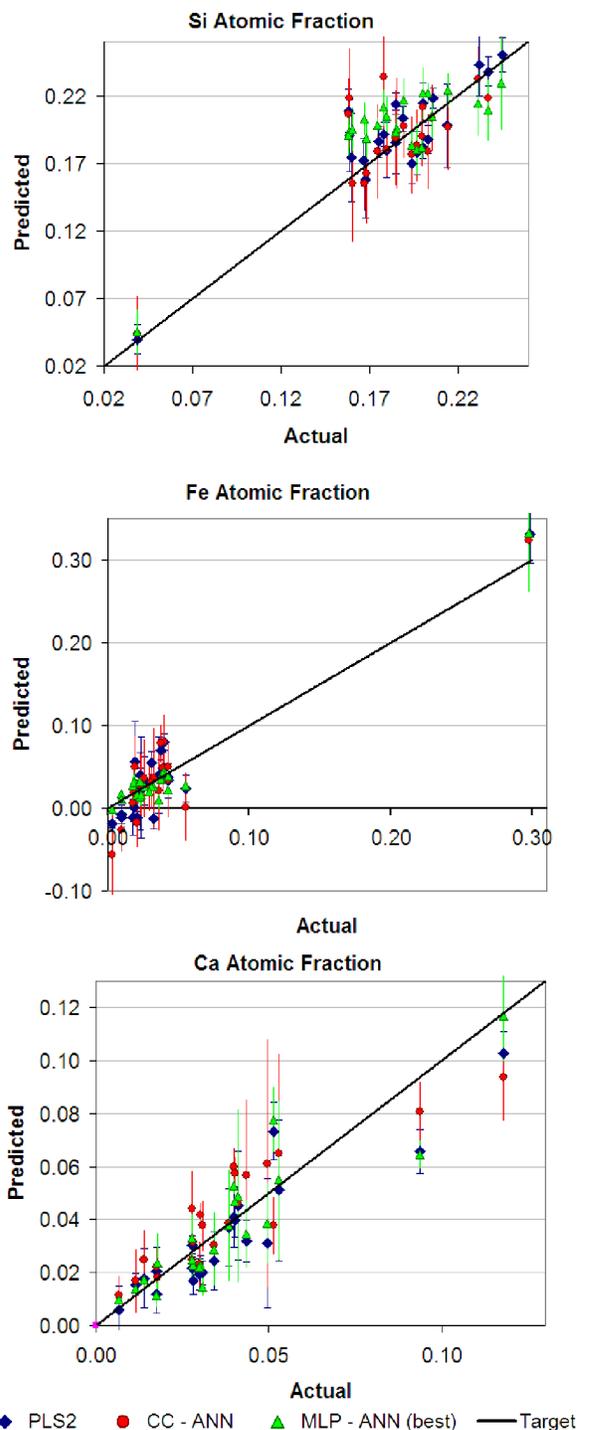
To conduct ANN calculations, we used the open-source Fast Artificial Neural Network (FANN) library [11] and the associated graphical interface, FannTool.

**Results:** Figure 1 shows the results of PLS2 predictions (blue diamonds), CC-ANN predictions (red circles) and MLP-ANN predictions (green triangles) of the atomic fraction of Si, Fe and Ca in half of the geologic samples. Each point on the plot is the average of five predictions corresponding to the five LIBS spots analyzed per sample. Error bars show the standard deviation of the predictions for the five points. A perfect prediction would fall along the black line.

Our preliminary results show that PLS2, CC-ANNs and MLP-ANNs perform comparably in predicting atomic fractions of major elements. Most samples were predicted well, as long as their composition was within the range of training set compositions. For Fe, a MLP-ANN with 5 hidden nodes performed better than PLS2 or ANN, as shown by the tight cluster of green points near the target line in figure 1b.

**Future work:** We are in the process of collecting LIBS spectra for ~200 geologic samples for which we have independent major and minor element abundances by standard XRF techniques in order to compile a comprehensive spectral database for use during the MSL mission. We will continue to compare MLP- and CC-ANNs and PLS to determine which method is best suited for quantitative ChemCam analyses. All three methods perform best when the training set is as similar as possible to the samples being predicted. Therefore we also intend to study clustering methods which would group similar spectra prior to generating a quantitative model.

**References:** [1] S. Maurice et al.(2005), *LPSC XXXVI*, Abstract #1735. [2] Clegg S.M. et al.(2008) *Spectrochimica Acta B*, 64, 79–88 [3] Labbe N. et al. (2008) *Applied Optics*, 47, 31. [4] Sirven J.B. et al. (2006) *Anal. Chem.*, 78(5), 1462-1469 [5] Inakollu P. et al., (2009) *Spectrochimica Acta B*, 64, 99–104 [6] Motto-Ros, V. et al., (2008) *JEOS - Rapid Publications* 3, 08011 [7] Morris, R.V. et al. *JGR*, Vol. 105, E1, 1757-1817, 2000 [8] Govindaraju, K. (1994) *Geostan-*



**Figure 1:** Predicted vs. actual atomic fractions for Si, Fe and Ca in the geologic samples, using PLS2 (blue), CC-ANN (red) and MLP-ANN (green).

*dards Newsletter*, Vol 18. [9] <http://www.camo.com/> [10] Geladi P. & Kowalski B.R. (1986) *Analytica Chimica Acta*, 186, 1-17 [11] Fahlman, S. E. & Lebiere, C. (1991) School of Comp. Sci., Carnegie Mellon Univ. [12] <http://leenissen.dk/fann/>